

# HIGH ORDER NUMERICAL SCHEMES FOR SECOND-ORDER FBSDES WITH APPLICATIONS TO STOCHASTIC OPTIMAL CONTROL \*

TAO KONG<sup>†</sup>, WEIDONG ZHAO<sup>‡</sup>, AND TAO ZHOU<sup>§</sup>

**Abstract.** This is one of our series papers on multistep schemes for solving forward backward stochastic differential equations (FBSDEs) and related problems. Here we extend (with non-trivial updates) our multistep schemes in [W. Zhao, Y. Fu and T. Zhou, SIAM J. Sci. Comput., 36 (2014), pp. A1731-A1751.] to solve the second order FBSDEs (2FBSDEs). The key feature of the multistep schemes is that the Euler method is used to discretize the forward SDE, which dramatically reduces the entire computational complexity. Moreover, it is shown that the usual quantities of interest (e.g., the solution tuple  $(Y_t, Z_t, A_t, \Gamma_t)$  in the 2FBSDEs) are still of high order accuracy. Several numerical examples are given to show the effective of the proposed numerical schemes. Applications of our numerical schemes for stochastic optimal control problems are also presented.

**Key words.** Multistep schemes, the Euler method, second-order forward backward stochastic differential equations, stochastic optimal control

**AMS subject classifications.** 60H35, 65H20, 65H30

**1. Introduction.** This work is concerned with numerical methods for the following coupled second order forward backward stochastic differential equations (2FBSDEs) which is defined on the filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, P)$  :

$$(1.1) \quad \begin{cases} X_t = x + \int_0^t b(s, \Theta_s) ds + \int_0^t \sigma(s, \Theta_s) dW_s, \\ Y_t = g(X_T) + \int_t^T f(s, \Theta_s) ds - \int_t^T Z_s dW_s, \quad s \in [0, T], \\ Z_s = Z_0 + \int_0^s A_s ds + \int_0^s \Gamma_s dW_s, \end{cases}$$

where  $\Theta_t = (X_t, Y_t, Z_t, A_t, \Gamma_t) \in \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d$  is the unknown,  $(\Omega, \mathcal{F}, P)$  is the given probability space,  $T > 0$  is the deterministic terminal time,  $\{W_t\}_{t \in [0, T]}$  is a  $d$ -dimensional Brownian Motion defined on  $(\Omega, \mathcal{F}, P)$  with the natural filtration  $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$  and all  $P$ -null sets in  $\mathcal{F}_0$ ,  $x \in \mathcal{F}_0$  is the initial condition of the forward SDE,  $\mathcal{S}^d$  is the set of all  $d \times d$  real-valued symmetric matrices, and

$$b : [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d \rightarrow \mathbb{R}^m, \quad \sigma : [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d \rightarrow \mathbb{R}^{m \times d}$$

are referred to the drift and diffusion coefficients of the forward SDE, respectively. While

$$f : [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d \rightarrow \mathbb{R}, \quad g : \mathbb{R}^m \rightarrow \mathbb{R}$$

---

\*This work is partially supported by the National Natural Science Foundations of China under grant numbers 91130003, 11201461 and 11171189, and Natural Science Foundation of Shandong Province under grant number ZR2011AZ002.

<sup>†</sup>School of Mathematics & Finance Institute, Shandong University, Jinan 250100, China. Email: vision.kt@gmail.com.

<sup>‡</sup>School of Mathematics & Finance Institute, Shandong University, Jinan 250100, China. Email: wdzhao@sdu.edu.cn.

<sup>§</sup>LSEC, Institute of Computational Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Email: tzhou@lsec.cc.ac.cn.

are referred to the generator and the terminal condition of the backward SDE, respectively. The two stochastic integrals with respect to the Brownian Motion  $\{W_t\}_{t \in [0, T]}$  are of the Itô type. A 5-tuple  $(X_t, Y_t, Z_t, \Gamma_t, A_t)$  is called an  $L^2$ -adapted solution of the 2FBSDEs (1.1) if it is  $\mathcal{F}_t$ -adapted, square integrable, and satisfies (1.1). Moreover, the 2FBSDE (1.1) is called *decoupled* if  $b$  and  $\sigma$  are independent of  $Y_t, Z_t, A_t$  and  $\Gamma_t$ .

The 2FBSDEs (1.1), as an extension of the BSDEs in the linear [2] or nonlinear cases [19], was first introduced by the cornerstone work of Cheridito, Soner, Touzi and Victoir [5], with a slightly different (yet equivalent) formula, and it was further investigated by Soner, Touzi and Zhang in [22]. The main motivation there is to give a precise connection between the 2FBSDEs and fully non-linear PDEs, in particular the Hamilton-Jacobi-Bellman equations and the Bellman-Isaacs equations which are widely used in stochastic control and in stochastic differential games. Such a connection leads to interesting stochastic representation results for fully nonlinear PDEs, generalizing the original (nonlinear) Feynman-Kac representations of linear and semi-linear parabolic PDEs, see e.g. [11, 16, 20, 21] and references therein.

From the numerical point of view, one can adopt these connections between PDEs and FBSDEs to design the so called probabilistic numerical methods for PDEs, by solving the equivalent FBSDEs (or 2FBSDEs). While there are a lot of work dealing with numerical schemes for BSDEs [1, 3, 4, 6, 13, 24, 25, 27], however, there are only a few work on numerical methods for FBSDEs [7, 9, 14, 15, 17, 23, 26, 28] and fully non-linear PDEs [8, 10]. Some of the above works are designed with high order accuracy that can however only be used to deal with low dimensional FBSDEs. While some of them are low order numerical methods that are suitable for solving high dimensional problems. In particular, we mention the work [10], where a numerical example for a 12-dimensional coupled FBSDE is reported, and it is shown by numerical test that the numerical method converges with order 1. Also, in [9], multistep schemes were proposed to solve multi-dimensional FBSDEs by using the sparse grid interpolation technique, and several multi-dimensional examples with dimension up to 6 are presented, and high order convergence rates up to 3 were obtained.

To the best of our knowledge, there is no related studies for high order numerical methods for 2FBSDEs. The main purpose in this work is to extend our multistep schemes in [26] (which is original designed for solving FBSDEs) to the use of solving 2FBSDEs. The key feature of the multistep schemes is that the Euler method is used to discrete the forward SDE, which dramatically reduces the entire computational complexity. Moreover, it is shown that the quantities of interest (e.g., the solution tuple  $(Y_t, Z_t, \Gamma_t, A_t)$  in the 2FBSDEs) are still of high order accuracy. Several numerical examples are given to show the effective of the proposed numerical schemes. With particular interests, applications of the proposed numerical schemes to stochastic optimal control problems are also presented and corresponding numerical examples are also given. It is worth to note that the multistep schemes proposed in this paper can also be used to solve a large class of fully nonlinear PDEs and we left this for our future studies.

The rest of the paper is organized as follows. In Section 2, we present some preliminaries. Then in Section 3, We shall discuss the multistep schemes for solving decoupled 2FBSDEs, and this is followed by extensions to coupled 2FBSDEs in Section 4. In Section 5, we shall present several numerical examples, and applications of our numerical methods for stochastic optimal control problems will also be presented. We finally give some concluding remarks in Section 6.

**2. Preliminaries.** We first introduce some notations that will be used in this paper. For  $x \in \mathbb{R}^d$ ,  $B \in \mathcal{S}^d$ , we set  $|x| := \sqrt{x_1^2 + \cdots + x_d^2}$ ,  $|B| := \sqrt{\sum_{i,j=1}^d B_{ij}^2}$ . We denote by  $C_b^k$  the set of functions  $g(x) : \mathbb{R} \rightarrow \mathbb{R}$  with uniformly bounded derivatives up to order  $k$ , and by  $C^{k_1, k_2}$  the set of functions  $f(t, x) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^m$  with continuous partial derivatives up to order  $k_1$  w.r.t.  $t \in \mathbb{R}$  and up to order  $k_2$  w.r.t.  $x \in \mathbb{R}^m$ .

**2.1. Diffusion process and its generator.** A stochastic process  $X_t$  is called a diffusion process starting at  $x_0$  and at the time  $t_0$  if it satisfies the following SDE

$$(2.1) \quad X_t = x_0 + \int_{t_0}^t b(s, X_s) ds + \int_{t_0}^t \sigma(s, X_s) dW_s, \quad t \in [t_0, T],$$

where  $b_s = b(s, X_s)$  and  $\sigma_s = \sigma(s, X_s)$  are measurable functions satisfying

$$(2.2) \quad \begin{aligned} |b(s, x)| + |\sigma(s, x)| &\leq C(1 + |x|), \quad x \in \mathbb{R}^m, s \in [t_0, T], \\ |b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| &\leq L|x - y|, \quad x, y \in \mathbb{R}^m, s \in [t_0, T]. \end{aligned}$$

It is well known that under conditions (2.2), the SDE (2.1) admits a unique solution. It is worth to note that, by the Markov property of the diffusion process, we have  $\mathbb{E}_t^x [X_s] = \mathbb{E} [X_s | X_t = x], \forall t \leq s$ .

Let  $X_t$  be the solution of (2.1). Then, for any given measurable function  $g : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $g(t, X_t)$  becomes a stochastic process. Moreover, we give the following definition

**DEFINITION 2.1.** Let  $X_s$  be a diffusion process in  $\mathbb{R}^m$  that satisfies (2.1). The generator  $A$  of  $X_s$  on  $g$  is defined by

$$(2.3) \quad Ag(t, x) = \lim_{s \downarrow t} \frac{\mathbb{E}_t^x [g(s, X_s)] - g(t, x)}{s - t}, \quad x \in \mathbb{R}^n.$$

Concerning the generator  $A$ , the following result holds [18]:

**THEOREM 1.** Let  $X_s$  be the diffusion process defined by the SDE (2.1). If  $f \in C^{1,2}([0, T] \times \mathbb{R}^m)$ , then we have

$$(2.4) \quad Af(t, x) = \mathcal{L}f(t, x), \quad Af(t, X_t) = \mathcal{L}f(t, X_t),$$

where the operator  $\mathcal{L}$  is defined by

$$(2.5) \quad \mathcal{L}\phi(t, x) := \phi_t(t, x) + \nabla_x \phi(t, x) b(t, x) + \frac{1}{2} \text{tr} (\sigma(t, x) \sigma^\top(t, x) \nabla_x^2 \phi(t, x)).$$

with  $\nabla_x \phi = (\partial_{x_1} \phi, \dots, \partial_{x_m} \phi)$ , and  $\nabla_x^2 \phi$  being the Hessian matrix of  $\phi$  with respect to the spatial variable  $x$ .

Note that  $Af(t, X_t) \in \mathcal{F}_t$  is a stochastic process. Furthermore, by using together the Itô's formula, Theorem 1 and the tower rule of conditional expectations, we have the following theorem.

**THEOREM 2.** Let  $t_0 < t$  be a fixed time, and  $x_0 \in \mathbb{R}^m$  be a fixed space point. If  $f \in C^{1,2}([0, T] \times \mathbb{R}^m)$  and  $\mathbb{E}_{t_0}^{x_0} [|Af(t, X_t)|] < +\infty$ , we have

$$(2.6) \quad \frac{d\mathbb{E}_{t_0}^{x_0} [f(t, X_t)]}{dt} = \mathbb{E}_{t_0}^{x_0} [Af(t, X_t)], \quad t \geq t_0.$$

Moreover, the following identity holds

$$(2.7) \quad \left. \frac{d\mathbb{E}_{t_0}^{x_0} [f(t, X_t)]}{dt} \right|_{t=t_0} = \left. \frac{d\mathbb{E}_{t_0}^{x_0} [f(t, \bar{X}_t)]}{dt} \right|_{t=t_0},$$

where  $\bar{X}_t$  is a diffusion process satisfying

$$(2.8) \quad \bar{X}_t = x_0 + \int_{t_0}^t \bar{b}_s ds + \int_{t_0}^t \bar{\sigma}_s dW_s,$$

with  $\bar{b} : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m, \bar{\sigma} : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$  being smooth functions satisfying

$$\bar{b}(t_0, x_0) = b(t_0, x_0), \quad \bar{\sigma}(t_0, x_0) = \sigma(t_0, x_0).$$

Note that by choosing different  $\bar{b}$  and  $\bar{\sigma}$ , the identity (2.7) yields different ways for approximating  $d\mathbb{E}_{t_0}^{x_0} [f(t, X_t)]/dt|_{t=t_0}$ . The computational complexity can be significantly reduced if appropriate choices are made.

**2.2. Solution regularity and representation theory of 2FBSDEs.** We now consider the following decoupled 2FBSDEs

$$(2.9) \quad \begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \\ -dY_t = f(t, X_t, Y_t, Z_t, \Gamma_t)dt - Z_t dW_t, \\ dZ_t = A_t dt + \Gamma_t dW_t \end{cases}$$

with a given terminal condition  $Y_T = g(X_T)$ . Under some standard assumptions (for details, please refer to [5, 22]), the following results are shown [5]:

**THEOREM 2.2.** *Let  $u = u(t, x)$  be the solution of the following fully nonlinear PDE*

$$(2.10) \quad \mathcal{L}u + f(t, x, u, \nabla_x u \sigma, \nabla_x(\nabla_x u \sigma)\sigma) = 0$$

with the terminal condition  $u(T, x) = g(x)$ , and let  $(X_t, Y_t, Z_t, \Gamma_t, A_t)$  be the solution of the 2FBSDE (2.9). Then we have

$$\begin{aligned} Y_t &= u(t, X_t), \quad Z_t = (\nabla_x u \sigma)(t, X_t), \\ \Gamma_t &= (\nabla_x(\nabla_x u \sigma)\sigma)(t, X_t), \quad A_t = (\mathcal{L}(\nabla_x u \sigma))(t, X_t). \end{aligned}$$

where the associate operator  $\mathcal{L}$  is defined by (2.5).

The above theorem provides a stochastic representation for solutions of fully nonlinear parabolic PDEs, generalizing the pioneer work on Feynman-Kac representations of linear and semi-linear parabolic PDEs [11, 16, 20, 21].

**2.3. Derivative approximation.** We now recall some basic results for numerical approximation of derivatives, and these results will play an important role for designing our high order numerical methods for 2FBSDEs.

Let  $u(t) \in C_b^{k+1}$  and  $t_i \in \mathbb{R}$  ( $0 \leq i \leq k$ ) satisfying  $t_0 < t_1 < \dots < t_k$ , where  $k$  is a positive integer. Let  $\Delta t_{0,i} = t_i - t_0, i = 0, 1, \dots, k$ . Then by Taylor's expansion, for each  $t_i, i = 0, 1, \dots, k$ , we have

$$u(t_i) = \sum_{j=0}^k \frac{(\Delta t_{0,i})^j}{j!} \frac{d^j u}{dt^j}(t_0) + O(\Delta t_{0,i})^{k+1}.$$

Then, we can deduce

$$\sum_{i=0}^k \alpha_{k,i} u(t_i) = \sum_{j=0}^k \frac{\sum_{i=0}^k \alpha_{k,i} (\Delta t_{0,i})^j}{j!} \frac{d^j u}{dt^j}(t_0) + O\left(\sum_{i=0}^k \alpha_{k,i} (\Delta t_{0,i})^{k+1}\right),$$

where  $\alpha_{k,i}$ ,  $i = 0, 1, \dots, k$ , are real nubmers. By choosing  $\alpha_{k,i}$  ( $i = 0, 1, \dots, k$ ) as

$$(2.11) \quad \frac{\sum_{i=0}^k \alpha_{k,i} (\Delta t_{0,i})^j}{j!} = \begin{cases} 1, & j = 1 \\ 0, & j \neq 1 \end{cases}.$$

One obtains a high order approximation

$$(2.12) \quad \frac{du}{dt}(t_0) = \sum_{i=0}^k \alpha_{k,i} u(t_i) + R_D,$$

where  $R_D = O(\sum_{i=0}^k \alpha_{k,i} (\Delta t_{0,i})^{k+1})$ . In particular, when  $\Delta t_{0,i} = i\Delta t$ , we get from (2.11) the following linear system for  $\alpha_{k,i}\Delta t$ ,

$$(2.13) \quad \sum_{i=1}^k i^j [\alpha_{k,i}\Delta t] = \begin{cases} 1, & j = 1 \\ 0, & j \neq 1 \end{cases}.$$

Note that the above system can be solved easily. We list  $\alpha_{k,i}\Delta t$  ( $i = 0, 1, \dots, k$ ) of the system (2.13) for  $k = 1, 2, \dots, 6$  in the following table.

TABLE 2.1

$\alpha_{k,i}\Delta t$	$i = 0$	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$
$k = 1$	-1	1					
$k = 2$	$-\frac{3}{2}$	2	$-\frac{1}{2}$				
$k = 3$	$-\frac{11}{6}$	3	$-\frac{3}{2}$	$\frac{1}{3}$			
$k = 4$	$-\frac{25}{12}$	4	-3	$\frac{4}{3}$	$-\frac{1}{4}$		
$k = 5$	$-\frac{137}{60}$	5	-5	$\frac{10}{3}$	$-\frac{5}{4}$	$\frac{1}{5}$	
$k = 6$	$-\frac{49}{20}$	6	$-\frac{15}{2}$	$\frac{20}{3}$	$-\frac{15}{4}$	$\frac{6}{5}$	$-\frac{1}{6}$

The above approximation schemes are very popular in numerical ODE community, and it is known that such an approximation scheme is unstable for  $k \geq 7$ , and this is why we have only listed the values of  $\alpha_{k,i}\Delta t$  for  $1 \leq k \leq 6$  in Table 2.1. For more details, one can refer to [26].

**3. Multistep schemes for decoupled 2FBSDEs.** In this section, we first consider the multistep schemes for solving decoupled 2FBSDEs (2.9), i.e., the forward SDE is independent of  $(Y_t, Z_t, A_t, \Gamma_t)$ . To begin, let  $N$  be a positive integer. For the time interval  $[t_0, T]$ , we introduce a regular time partition:

$$t_0 < t_1 < \dots < t_N = T.$$

We set  $\Delta t_{t_n,k} = t_{n+k} - t_n$  for  $n, \in \{1, 2, \dots, N\}$  and  $k \in \mathbb{N}$  satisfying  $n+k \leq N$ . For  $t \geq t_n$ , we denote  $\Delta W_{t_n,k} = W_{t_{n+k}} - W_{t_n}$ ,  $\Delta t_{t_n,t} = t - t_n$  and  $\Delta W_{t_n,t} = W_t - W_{t_n}$ ,

**3.1. Four reference ODEs.** Let  $\Theta_t = (X_t, Y_t, Z_t, A_t, \Gamma_t)$  be the solution of the decoupled FBSDEs (2.9) with terminal condition  $g(X_T)$ . We set  $\mathbb{E}_{t_n}^x[\cdot] = \mathbb{E}[\cdot | \mathcal{F}_{t_n}^{t_n, x}]$ . By taking conditional expectation  $\mathbb{E}_{t_n}^x[\cdot]$  on both sides of the last two equations in (2.9), we obtain

$$(3.1) \quad \mathbb{E}_{t_n}^x[Y_t] = \mathbb{E}_{t_n}^x[g(X_T)] + \int_t^T \mathbb{E}_{t_n}^x[f(s, \Theta_s)] ds, \quad t \in [t_n, T],$$

$$(3.2) \quad \mathbb{E}_{t_n}^x[Z_t] = \mathbb{E}_{t_n}^x[Z_{t_n}] + \int_{t_n}^t \mathbb{E}_{t_n}^x[A_s] ds, \quad t \in [t_n, T],$$

By taking derivative with respect to  $t$  in (3.1)-(3.2), one gets the following two reference ODEs:

$$(3.3) \quad \frac{d\mathbb{E}_{t_n}^x[Y_t]}{dt} = -\mathbb{E}_{t_n}^x[f(t, \Theta_t)], \quad t \in [t_n, T],$$

$$(3.4) \quad \frac{d\mathbb{E}_{t_n}^x[Z_t]}{dt} = \mathbb{E}_{t_n}^x[A_t], \quad t \in [t_n, T]$$

Under sufficient regularity assumptions on the given data, the integrand  $\mathbb{E}_{t_n}^x[f(s, \Theta_s)]$  and  $\mathbb{E}_{t_n}^x[A_t]$  are continuous at  $s = t$ . Note that we also have

$$Y_{t_n} = Y_t + \int_{t_n}^t f(s, \Theta_s) ds - \int_{t_n}^t Z_s dW_s, \quad t \in [t_n, T]$$

By multiplying both sides of the above equation by  $\Delta W_{t_n, t}^\top$  (where " $\top$ " stands for the transposition), and taking the conditional expectation  $\mathbb{E}_{t_n}^x[\cdot]$  on both sides of the derived equation, we obtain, for  $t \in [t_n, T]$

$$(3.5) \quad 0 = \mathbb{E}_{t_n}^x[Y_t \Delta W_{t_n, t}^\top] + \int_{t_n}^t \mathbb{E}_{t_n}^x[f(s, \Theta_s) \Delta W_{t_n, s}^\top] ds - \int_{t_n}^t \mathbb{E}_{t_n}^x[Z_s] ds,$$

Similarly, for the last equation in (2.9), we have for  $t \in [t_n, T]$

$$(3.6) \quad 0 = \mathbb{E}_{t_n}^x[Z_t^\top \Delta W_{t_n, t}^\top] - \int_{t_n}^t \mathbb{E}_{t_n}^x[A_s^\top \Delta W_{t_n, s}^\top] ds - \int_{t_n}^t \mathbb{E}_{t_n}^x[\Gamma_s] ds.$$

Now, by taking derivative with respect to  $t \in [t_n, T]$  on both sides, one gets the following reference ODEs:

$$(3.7) \quad \frac{d\mathbb{E}_{t_n}^x[Y_t \Delta W_{t_n, t}^\top]}{dt} = -\mathbb{E}_{t_n}^x[f(t, \Theta_t) \Delta W_{t_n, t}^\top] + \mathbb{E}_{t_n}^x[Z_t], \quad t \in [t_n, T]$$

$$(3.8) \quad \frac{d\mathbb{E}_{t_n}^x[Z_t^\top \Delta W_{t_n, t}^\top]}{dt} = \mathbb{E}_{t_n}^x[A_t^\top \Delta W_{t_n, t}^\top] + \mathbb{E}_{t_n}^x[\Gamma_t], \quad t \in [t_n, T]$$

The equations (3.3), (3.4), (3.7), and (3.8) are reference ODEs for the decoupled 2FBSDEs (2.9). Our multistep numerical schemes will be constructed by approximating the associate derivatives and the conditional expectations in these ODEs.

**3.2. The semi-discrete scheme.** Motivated by Theorem 2, we choose smooth functions  $\bar{b}(t, x)$  and  $\bar{\sigma}(t, x)$  for  $t \in [t_n, T]$  and  $x \in \mathbb{R}^m$  with the constraints  $\bar{b}(t_n, x) = b(t_n, x)$  and  $\bar{\sigma}(t_n, x) = \sigma(t_n, x)$ . Let  $\bar{X}_t^{t_n, x}$  be the diffusion process defined by

$$(3.9) \quad \bar{X}_t^{t_n, x} = x + \int_{t_n}^t \bar{b}(s, \bar{X}_s^{t_n, x}) ds + \int_{t_n}^t \bar{\sigma}(s, \bar{X}_s^{t_n, x}) dW_s.$$

Note that  $Y_t, Z_t, \Gamma_t, A_t$  are all functions of  $(t, X_t)$ . Now, let  $(\bar{Y}_t^{t_n, x}, \bar{Z}_t^{t_n, x})$  be the value of function  $(Y_t, Z_t)$  at the time-space point  $(t, \bar{X}_t^{t_n, x})$ . , by Theorem 2, we have

$$\begin{aligned} \left. \frac{d\mathbb{E}_{t_n}^x [Y_t]}{dt} \right|_{t=t_n} &= \left. \frac{d\mathbb{E}_{t_n}^x [\bar{Y}_t^{t_n, x}]}{dt} \right|_{t=t_n}, & \left. \frac{d\mathbb{E}_{t_n}^x [Y_t \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n} &= \left. \frac{d\mathbb{E}_{t_n}^x [\bar{Y}_t^{t_n, x} \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n}, \\ \left. \frac{d\mathbb{E}_{t_n}^x [Z_t]}{dt} \right|_{t=t_n} &= \left. \frac{d\mathbb{E}_{t_n}^x [\bar{Z}_t^{t_n, x}]}{dt} \right|_{t=t_n}, & \left. \frac{d\mathbb{E}_{t_n}^x [Z_t^\top \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n} &= \left. \frac{d\mathbb{E}_{t_n}^x [(\bar{Z}_t^{t_n, x})^\top \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n}. \end{aligned}$$

Now applying (2.12) to  $\left. \frac{d\mathbb{E}_{t_n}^x [\bar{Y}_t^{t_n, x}]}{dt} \right|_{t=t_n}$  and  $\left. \frac{d\mathbb{E}_{t_n}^x [\bar{Y}_t^{t_n, x} \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n}$ ,  $\left. \frac{d\mathbb{E}_{t_n}^x [\bar{Z}_t^{t_n, x}]}{dt} \right|_{t=t_n}$  and  $\left. \frac{d\mathbb{E}_{t_n}^x [(\bar{Z}_t^{t_n, x})^\top \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n}$ , we deduce

$$(3.10) \quad \left\{ \begin{aligned} \left. \frac{d\mathbb{E}_{t_n}^x [Y_t]}{dt} \right|_{t=t_n} &= \sum_{i=0}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Y}_{t_{n+i}}^{t_n, x}] + \bar{R}_{y,n}^k, \\ \left. \frac{d\mathbb{E}_{t_n}^x [Y_t \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n} &= \sum_{i=1}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Y}_{t_{n+i}}^{t_n, x} \Delta W_{n,i}^\top] + \bar{R}_{z,n}^k, \\ \left. \frac{d\mathbb{E}_{t_n}^x [Z_t]}{dt} \right|_{t=t_n} &= \sum_{i=0}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Z}_{t_{n+i}}^{t_n, x}] + \bar{R}_{A,n}^k, \\ \left. \frac{d\mathbb{E}_{t_n}^x [Z_t^\top \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n} &= \sum_{i=1}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [(\bar{Z}_{t_{n+i}}^{t_n, x})^\top \Delta W_{n,i}^\top] + \bar{R}_{\Gamma,n}^k, \end{aligned} \right.$$

Here  $\alpha_{k,i}$  are given by (2.11), and  $\bar{R}_{y,n}^k, \bar{R}_{z,n}^k, \bar{R}_{A,n}^k$  and  $\bar{R}_{\Gamma,n}^k$  are the associate truncation errors, i.e.

$$\begin{aligned} \bar{R}_{y,n}^k &= \left. \frac{d\mathbb{E}_{t_n}^x [Y_t]}{dt} \right|_{t=t_n} - \sum_{i=0}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Y}_{t_{n+i}}^{t_n, x}], \\ \bar{R}_{z,n}^k &= \left. \frac{d\mathbb{E}_{t_n}^x [Y_t \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n} - \sum_{i=1}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Y}_{t_{n+i}}^{t_n, x} \Delta W_{n,i}^\top], \\ \bar{R}_{A,n}^k &= \left. \frac{d\mathbb{E}_{t_n}^x [Z_t]}{dt} \right|_{t=t_n} - \sum_{i=0}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Z}_{t_{n+i}}^{t_n, x}], \\ \bar{R}_{\Gamma,n}^k &= \left. \frac{d\mathbb{E}_{t_n}^x [Z_t^\top \Delta W_{t_n, t}^\top]}{dt} \right|_{t=t_n} - \sum_{i=1}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [(\bar{Z}_{t_{n+i}}^{t_n, x})^\top \Delta W_{n,i}^\top]. \end{aligned}$$

By inserting (3.10) into (3.3), (3.4), (3.7) and (3.8), respectively, we deduce

$$(3.11) \quad \begin{cases} \sum_{i=0}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Y}_{t_{n+i}}] = -f(t_n, x, Y_{t_n}, Z_{t_n}) + R_{y,n}^k, \\ \sum_{i=1}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Y}_{t_{n+i}} \Delta W_{n,i}^\top] = Z_{t_n} + R_{z,n}^k, \\ \sum_{i=0}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Z}_{t_{n+i}}] = A_{t_n} + R_{A,n}^k, \\ \sum_{i=1}^k \alpha_{k,i} \mathbb{E}_{t_n}^x [\bar{Z}_{t_{n+i}}^\top \Delta W_{n,i}^\top] = \Gamma_{t_n} + R_{\Gamma,n}^k, \end{cases}$$

where  $R_{y,n}^k = -\bar{R}_{y,n}^k$ ,  $R_{z,n}^k = -\bar{R}_{z,n}^k$ ,  $R_{\Gamma,n}^k = -\bar{R}_{\Gamma,n}^k$  and  $R_{A,n}^k = -\bar{R}_{A,n}^k$ .

Let  $Y^n$ ,  $Z^n$ ,  $A^n$  and  $\Gamma^n$  be four random variables that represent the approximate values of the solutions  $Y_t$ ,  $Z_t$ ,  $A_t$  and  $\Gamma_t$  of the 2FBSDE in (2.9) at time  $t_n$ , respectively. By removing the truncation error terms  $R_{y,n}^k$ ,  $R_{z,n}^k$ ,  $R_{A,n}^k$  and  $R_{\Gamma,n}^k$  from (3.11), we obtain the following semi-discrete numerical scheme for solving the 2FBSDEs (2.9):

SCHEME 1. Assume that random variables  $Y^{N-i}$  and  $Z^{N-i}$ ,  $i = 0, 1, \dots, k-1$  are known. For  $n = N-k, \dots, 0$ , with  $X_t^{t_n, X^n}$  being the solution of (3.9), solve  $Y^n = Y^n(X^n)$ ,  $Z^n = Z^n(X^n)$ ,  $A^n = A^n(X^n)$  and  $\Gamma^n = \Gamma^n(X^n)$  by

$$(3.12) \quad Z^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{X^n} [\bar{Y}^{n+j} \Delta W_{n,j}^\top],$$

$$(3.13) \quad \Gamma^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{X^n} [(\bar{Z}^{n+j})^\top \Delta W_{n,j}^\top],$$

$$(3.14) \quad A^n = \sum_{j=0}^k \alpha_{k,j} \mathbb{E}_{t_n}^{X^n} [\bar{Z}^{n+j}],$$

$$(3.15) \quad -\alpha_{k,0} Y^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{X^n} [\bar{Y}^{n+j}] + f(t_n, X^n, Y^n, Z^n, \Gamma^n),$$

where  $\bar{Y}^{n+j}$  and  $\bar{Z}^{n+j}$  are the values of  $Y^{n+j}$  and  $Z^{n+j}$  at the space points  $X_{t_{n+j}}^{t_n, X^n}$ .

In addition, we choose  $\bar{b}$  and  $\bar{\sigma}$  in (3.9) as

$$(3.16) \quad \bar{b}(s, X_s^{t_n, x}) = b(t_n, x), \quad \bar{\sigma}(s, X_s^{t_n, x}) = \sigma(t_n, x), \quad \forall s \in [t_n, T].$$

Then, (3.9) yields the Euler discretization scheme for the forward SDE. In this case, Scheme 1 becomes

SCHEME 2. Assume that random variables  $Y^{N-i}$  and  $Z^{N-i}$ ,  $i = 0, 1, \dots, k-1$ , are known. For  $n = N-k, \dots, 0$ , solve the random variables  $X^{n,j}$ , ( $j = 1, 2, \dots, k$ ),



$Y^n$ ,  $Z^n$ ,  $A^n$  and  $\Gamma^n$  by

$$(3.17) \quad X^{n,j} = X^n + b(t_n, X^n) \Delta t_{n,j} + \sigma(t_n, X^n) \Delta W_{n,j}, \quad j = 1, \dots, k,$$

$$(3.18) \quad Z^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{X^n} [\bar{Y}^{n+j} \Delta W_{n,j}^\top],$$

$$(3.19) \quad A^n = \sum_{j=0}^k \alpha_{k,j} \mathbb{E}_{t_n}^{X^n} [\bar{Z}^{n+j}],$$

$$(3.20) \quad \Gamma^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{X^n} [(\bar{Z}^{n+j})^\top \Delta W_{n,j}^\top],$$

$$(3.21) \quad -\alpha_{k,0} Y^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{X^n} [\bar{Y}^{n+j}] + f(t_n, X^n, Y^n, Z^n, \Gamma^n),$$

where  $\bar{Y}^{n+j}$  and  $\bar{Z}^{n+j}$  stand for the values of  $Y^{n+j}$  and  $Z^{n+j}$  at the space points  $X^{n,j}$ , respectively.

If the functions  $\mathbb{E}_{t_n}^x [Y_t]$ ,  $\mathbb{E}_{t_n}^x [Y_t \Delta W_{t_n,t}^\top]$ ,  $\mathbb{E}_{t_n}^x [Z_t]$ ,  $\mathbb{E}_{t_n}^x [Z_t^\top \Delta W_{t_n,t}^\top]$  and their derivatives (with respect to  $t$ ) up to order  $k+1$  are bounded, we have the estimates

$$(3.22) \quad \bar{R}_{y,n}^k = O(\Delta t)^k, \quad \bar{R}_{z,n}^k = O(\Delta t)^k, \quad \bar{R}_{A,n}^k = O(\Delta t)^k, \quad \bar{R}_{\Gamma,n}^k = O(\Delta t)^k.$$

**3.3. The fully discrete scheme.** We now propose our fully discrete schemes. To this end, we first introduce the time-space partition. Given a time partition  $\mathcal{T} := \{t_0, t_1, \dots, t_N\}$ , we introduce a series of space partitions  $\mathcal{D}_h := \{\mathcal{D}_{h_n}^n\}_{n=0,1,\dots,N}$  for space  $\mathbb{R}^m$ , with each  $\mathcal{D}_{h_n}^n$  corresponds to the time level  $t_n \in \mathcal{T}$  and  $\mathbf{h} = (h_0, h_1, \dots, h_N)$  being the partition density of each  $\mathcal{D}_{h_n}^n$ . We use  $\mathcal{D}_h^n$  for simplicity if the context is clear. In addition, If the space partitions w.r.t. all the time levels are the same, then  $\mathcal{D}_h$  is used to represent the unified space partition.

For the space partition  $\mathcal{D}_{h_n}^n = \{x_j \in \mathbb{R}^m | j \in I, I \subset \mathbb{N}\}$ ,  $x_j$  is called the grid points and  $I$  is called the index set of the grid. The partition density  $h_n$  of  $\mathcal{D}_{h_n}^n$  is defined by  $h_n = \max_{x_i, x_j \in \mathcal{D}_{h_n}^n} d(x_i, x_j)$ , with  $x_i, x_j$  adjacent in  $\mathcal{D}_{h_n}^n$ , where  $d(\cdot, \cdot)$  is the distance between two points in  $\mathbb{R}^m$ . Given  $x \in \mathbb{R}^m$ , a neighbor grid  $U_{n,x}^{\mathcal{D}}$  of  $x$ , is a finite-element subset of  $\mathcal{D}_{h_n}^n$  satisfying  $\min_{x_i \in U_{n,x}^{\mathcal{D}}} d(x, x_i) < \min_{x_i \in \mathcal{D}_{h_n}^n \setminus U_{n,x}^{\mathcal{D}}} d(x, x_i)$ .

With the above spacial partition, we seek to solving  $Y^n$ ,  $Z^n$ ,  $A^n$  and  $\Gamma^n$  at grid point  $x \in \mathcal{D}_h^n$ . More precisely, for each  $x \in \mathcal{D}_h^n, n = N-k, \dots, 0$ , we aim at solving  $Y^n$ ,  $Z^n$ ,  $\Gamma^n$  and  $A^n$  by

$$(3.23) \quad \left\{ \begin{array}{l} Z^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^x [\bar{Y}^{n+j} \Delta W_{n,j}^\top], \\ \Gamma^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^x [(\bar{Z}^{n+j})^\top \Delta W_{n,j}^\top], \\ A^n = \sum_{j=0}^k \alpha_{k,j} \mathbb{E}_{t_n}^x [\bar{Z}^{n+j}], \\ -\alpha_{k,0} Y^n = \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^x [\bar{Y}^{n+j}] + f(t_n, x, Y^n, Z^n, \Gamma^n), \end{array} \right.$$

where  $\bar{Y}^{n+j}$  and  $\bar{Z}^{n+j}$  are respectively the values of  $Y^{n+j}$  and  $Z^{n+j}$  at the space points  $\bar{X}^{n,j}$  that is defined by

$$(3.24) \quad \bar{X}^{n,j} = x + b(t_n, x)\Delta t_{n,j} + \sigma(t_n, x)\Delta W_{n,j}, \quad j = 1, \dots, k.$$

Plug (3.24) into the conditional expectation of (3.23), since  $\bar{Y}, \bar{Z}, \bar{\Gamma}$  and  $\bar{A}$  are all functions of  $X^{n,j}$ , we have

$$\begin{aligned} \mathbb{E}_{t_n}^x [\bar{Y}^{n+j} \Delta W_{n,j}^\top] &= \mathbb{E}_{t_n}^x [Y^{n+j}(x + b(t_n, x)\Delta t_{n,j} + \sigma(t_n, x)\Delta W_{n,j})\Delta W_{n,j}^\top] \\ \mathbb{E}_{t_n}^x [\bar{Z}^{n+j}] &= \mathbb{E}_{t_n}^x [Z_{t_{n+j}}(x + b(t_n, x)\Delta t_{n,j} + \sigma(t_n, x)\Delta W_{n,j})] \\ \mathbb{E}_{t_n}^x [(\bar{Z}^{n+j})^\top \Delta W_{n,j}^\top] &= \mathbb{E}_{t_n}^x [(Z^{n+j})^\top (x + b(t_n, x)\Delta t_{n,j} + \sigma(t_n, x)\Delta W_{n,j})\Delta W_{n,j}^\top] \\ \mathbb{E}_{t_n}^x [\bar{Y}^{n+j}] &= \mathbb{E}_{t_n}^x [Y^{n+j}(x + b(t_n, x)\Delta t_{n,j} + \sigma(t_n, x)\Delta W_{n,j})] \end{aligned}$$

Note that quadrature methods should be applied when approximating the conditional expectations above. Any efficient quadrature rules such as the Monte-Carlo methods, the quasi-Monte-Carlo methods, and the Gaussian quadrature methods can be used. Here, we will use the Gaussian quadratures based on the zeros of Hermite polynomials, for details, one can refer to [26]. In what follows, we shall denote by  $\hat{\mathbb{E}}^{n,x}[\cdot]$  the numerical quadrature operator for the conditional expectations.

Furthermore, when quadrature method is applied to approximate the conditional expectations, non-grid information might be needed. That's to say, for  $x \in \mathcal{D}_h^n$ , points ( $X^{n,j}$  defined by (3.24) e.t.c.) that are not in  $\mathcal{D}_h^{n+j}$  are needed when approximate the conditional expectation at time level  $t_{n+j}$ . Thus, interpolation methods are also needed. We shall denote by  $\mathbb{I}_{\mathcal{D}}$ , a local interpolation operator, such that  $\mathbb{I}_{\mathcal{D}}g$  is the continuous function interpolated by the values of  $g$  at the grid points  $\mathcal{D}$ . We also denote by  $\mathbb{I}_{\mathcal{D}}^n$  with the superscript  $n$  indicating the interpolation method at time level  $t_n$ . If the grid is a neighbor grid  $U_{n,x}^{\mathcal{D}}$  of  $x$ , we denote by  $\mathbb{I}_{\mathcal{D},x}^n$  instead of  $\mathbb{I}_{U_{n,x}^{\mathcal{D}}}^n$ , for simplicity. Note that any interpolation methods can be used here, however, care should be made if one wants to guarantee the stability and accuracy.

Now by introducing the operators  $\hat{\mathbb{E}}^{n,x}[\cdot]$  and  $\mathbb{I}_{\mathcal{D},x}^n$  (or more precisely  $\mathbb{I}_{\mathcal{D},\bar{X}_{t_{n+j}}^{n,x}}^{n+j}$ ) one can rewrite (3.11) in the following equivalent form

$$\begin{aligned} Z_{t_n} &= \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D},\bar{X}_{t_{n+j}}^{n,x}}^{n+j} Y_{t_{n+j}} \Delta W_{n,j}^\top \right] - R_{z,n}^k + R_{z,n}^{k,\mathbb{E}} + R_{z,n}^{k,\mathbb{I}}, \\ \Gamma_{t_n} &= \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D},\bar{X}_{t_{n+j}}^{n,x}}^{n+j} Z_{t_{n+j}}^\top \Delta W_{n,j}^\top \right] - R_{\Gamma,n}^k + R_{\Gamma,n}^{k,\mathbb{E}} + R_{\Gamma,n}^{k,\mathbb{I}}, \\ (3.25) \quad A_{t_n} &= \sum_{j=0}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D},\bar{X}_{t_{n+j}}^{n,x}}^{n+j} Z_{t_{n+j}} \right] - R_{A,n}^k + R_{A,n}^{k,\mathbb{E}} + R_{A,n}^{k,\mathbb{I}}, \\ -\alpha_{k,0} Y_{t_n} &= \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D},\bar{X}_{t_{n+j}}^{n,x}}^{n+j} Y_{t_{n+j}} \right] + f(t_n, x, Y_{t_n}, Z_{t_n}) \\ &\quad + R_{y,n}^k + R_{y,n}^{k,\mathbb{E}} + R_{y,n}^{k,\mathbb{I}}, \end{aligned}$$

where

$$\begin{aligned}
R_{z,n}^{k,\mathbb{E}} &= \sum_{j=1}^k \alpha_{k,j} (\mathbb{E}_{t_n}^x - \hat{\mathbb{E}}^{n,x}) [\bar{Y}_{t_{n+j}} \Delta W_{n,j}^\top], \\
R_{z,n}^{k,\mathbb{I}} &= \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{x,h} \left[ (\bar{Y}_{t_{n+j}} - \mathbb{I}_{\mathcal{D}, \bar{X}_{t_{n+j}}}^{n+j} Y_{t_{n+j}}) \Delta W_{n,j}^\top \right], \\
R_{A,n}^{k,\mathbb{E}} &= \sum_{j=1}^k \alpha_{k,j} (\mathbb{E}_{t_n}^x - \hat{\mathbb{E}}^{n,x}) [\bar{Z}_{t_{n+j}}], \\
R_{A,n}^{k,\mathbb{I}} &= \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{x,h} \left[ (\bar{Z}_{t_{n+j}} - \mathbb{I}_{\mathcal{D}, \bar{X}_{t_{n+j}}}^{n+j} Z_{t_{n+j}}) \right], \\
R_{\Gamma,n}^{k,\mathbb{E}} &= \sum_{j=1}^k \alpha_{k,j} (\mathbb{E}_{t_n}^x - \hat{\mathbb{E}}^{n,x}) [\bar{Z}_{t_{n+j}}^\top \Delta W_{n,j}^\top], \\
R_{\Gamma,n}^{k,\mathbb{I}} &= \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{x,h} \left[ (\bar{Z}_{t_{n+j}}^\top - \mathbb{I}_{\mathcal{D}, \bar{X}_{t_{n+j}}}^{n+j} Z_{t_{n+j}}^\top) \Delta W_{n,j}^\top \right], \\
R_{y,n}^{k,\mathbb{E}} &= - \sum_{j=1}^k \alpha_{k,j} (\mathbb{E}_{t_n}^x - \hat{\mathbb{E}}^{n,x}) [\bar{Y}^{n+j}], \\
R_{y,n}^{k,\mathbb{I}} &= - \sum_{j=1}^k \alpha_{k,j} \mathbb{E}_{t_n}^{x,h} \left[ \bar{Y}_{t_{n+j}} - \mathbb{I}_{\mathcal{D}, \bar{X}_{t_{n+j}}}^{n+j} Y_{t_{n+j}} \right].
\end{aligned}$$

The four terms  $R_{y,n}^{k,\mathbb{E}}$ ,  $R_{z,n}^{k,\mathbb{E}}$ ,  $R_{A,n}^{k,\mathbb{E}}$  and  $R_{\Gamma,n}^{k,\mathbb{E}}$  are the error terms resulted from approximating conditional expectations, and the other four terms  $R_{y,n}^{k,\mathbb{I}}$ ,  $R_{z,n}^{k,\mathbb{I}}$ ,  $R_{A,n}^{k,\mathbb{I}}$  and  $R_{\Gamma,n}^{k,\mathbb{I}}$  are the error terms caused by numerical interpolations.

By removing those error terms  $R_{y,n}^{k,\mathbb{E}}$ ,  $R_{y,n}^{k,\mathbb{I}}$ ,  $R_{y,n}^{k,\mathbb{E}}$ ,  $R_{z,n}^{k,\mathbb{E}}$ ,  $R_{z,n}^{k,\mathbb{I}}$ ,  $R_{A,n}^{k,\mathbb{E}}$ ,  $R_{A,n}^{k,\mathbb{I}}$ ,  $R_{\Gamma,n}^{k,\mathbb{E}}$  and  $R_{\Gamma,n}^{k,\mathbb{I}}$  from (3.25), we propose our fully discrete scheme for solving 2FBSDEs as follows:

SCHEME 3. Assume random variables  $Y^{N-i}$  and  $Z^{N-i}$  defined on  $\mathcal{D}_h^{N-i}$ ,  $i = 0, 1, \dots, k-1$ , are known. For  $n = N-k, \dots, 0$ , and for each  $x \in \mathcal{D}_h^n$ , solve  $X^n$ ,  $Y^n$ ,  $Z^n$ ,  $A^n$  and  $\Gamma^n$  by

$$(3.26) \quad X^{n,j} = x + b(t_n, x) \Delta t_{n,j} + \sigma(t_n, x) \Delta W_{n,j}, \quad j = 1, \dots, k,$$

$$(3.27) \quad Z^n = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}_{n,j}}^{n+j} Y^{n+j} \Delta W_{n,j}^\top \right],$$

$$(3.28) \quad \Gamma^n = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}_{n,j}}^{n+j} (Z^{n+j})^\top \Delta W_{n,j}^\top \right],$$

$$(3.29) \quad A^n = \sum_{j=0}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}_{n,j}}^{n+j} Z^{n+j} \right],$$

$$(3.30) \quad -\alpha_{k,0} Y^n = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}_{n,j}}^{n+j} Y^{n+j} \right] + f(t_n, x, Y^n, Z^n, \Gamma^n).$$

In Scheme 3, for a fixed integer  $k$ , to get the values  $Y^n$ ,  $Z^n$ ,  $A^n$  and  $\Gamma^n$  on each time level  $t_n$  at the grid  $\mathcal{D}_h^n$ , we shall solve the values at  $x \in \mathcal{D}_h^n$  one by one. Since they are independent with each other, the process can be completely parallel. For a fixed  $x \in \mathcal{D}_h^n$ , in Scheme 3, we firstly solve  $X^{n,j}$  by the Euler scheme (3.26) for  $1 \leq j \leq k$ ; and then solve  $Z^n$ ,  $\Gamma^n$  and  $A^n$  by (3.27), (3.28) and (3.29) explicitly; finally, solve  $Y^n$  by (3.30) implicitly. Thus some iteration methods are required for solving  $Y^n$ .

If the function  $f(t_n, x, y, z, \gamma)$  is Lipschitz continuous with respect to  $y$ , for small time partition step size  $\Delta t_n$ , we can use the following iteration procedure to solve  $Y^n$

$$(3.31) \quad \alpha_{k,0} Y^{n,l+1} = - \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} Y^{n+j} \right] - f(t_n, x, Y^{n,l}, Z^n, \Gamma^n),$$

with the iteration stopping condition  $|Y^{n,l+1} - Y^{n,l}| \leq \epsilon_0$ , where  $\epsilon_0 > 0$  is a given tolerance. In case the function  $f(t_n, x, y, z, \gamma)$  is differentiable with respect to  $y$ , to accelerate the convergence rate, the Newton iteration method can be applied, i.e.

$$Y^{n,l+1} = Y^{n,l} - \frac{\alpha_{k,0} Y^{n,l} + \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} Y^{n+j} \right] + f(t_n, x, Y^{n,l}, Z^n, \Gamma^n)}{\alpha_{k,0} + f_y(t_n, x, Y^{n,l}, Z^n, \Gamma^n)}$$

Note that the local truncation errors of Scheme 3 consist of twelve terms  $R_{y,n}^k$ ,  $R_{y,n}^{k,\mathbb{E}}$ ,  $R_{y,n}^{k,\mathbb{I}}$ ,  $R_{z,n}^k$ ,  $R_{z,n}^{k,\mathbb{E}}$ ,  $R_{z,n}^{k,\mathbb{I}}$ ,  $R_{A,n}^k$ ,  $R_{A,n}^{k,\mathbb{E}}$ ,  $R_{A,n}^{k,\mathbb{I}}$ ,  $R_{\Gamma,n}^k$ ,  $R_{\Gamma,n}^{k,\mathbb{E}}$  and  $R_{\Gamma,n}^{k,\mathbb{I}}$ . The four terms  $R_{y,n}^k$ ,  $R_{z,n}^k$ ,  $R_{A,n}^k$  and  $R_{\Gamma,n}^k$  defined respectively in (3.10) come from the approximations of the derivative, and the four terms  $R_{y,n}^{k,\mathbb{I}}$ ,  $R_{z,n}^{k,\mathbb{I}}$ ,  $R_{A,n}^{k,\mathbb{I}}$  and  $R_{\Gamma,n}^{k,\mathbb{I}}$  defined in (3.25) are the local interpolation errors. For these eight terms, when data  $b$ ,  $\sigma$ ,  $f$  and  $g$  are smooth, the following estimates hold (provided that degree  $r$  polynomials interpolation is used)

$$(3.32) \quad R_{y,n}^k, R_{z,n}^k, R_{A,n}^k, R_{\Gamma,n}^k = O(\Delta t)^k, \quad R_{y,n}^{k,\mathbb{I}}, R_{z,n}^{k,\mathbb{I}}, R_{A,n}^{k,\mathbb{I}}, R_{\Gamma,n}^{k,\mathbb{I}} = O(h^{r+1}).$$

The other four terms  $R_{y,n}^{k,\mathbb{E}}$ ,  $R_{z,n}^{k,\mathbb{E}}$ ,  $R_{A,n}^{k,\mathbb{E}}$  and  $R_{\Gamma,n}^{k,\mathbb{E}}$  are the local truncation errors resulted from the approximations of the conditional expectations. Noticed that these conditional expectations are functions of Gaussian random variables, thus, one can construct efficient Gauss-Hermite quadrature rule for their approximations.

**4. Extensions to coupled 2FBSDEs.** In this section we extend Scheme 3 to solve the fully coupled 2FBSDE. More precisely, we first update Scheme 3 into the following scheme

SCHEME 4. Assume random variables  $Y^{N-i}$  and  $Z^{N-i}$  defined on  $\mathcal{D}_h^{N-i}$ ,  $i = 0, 1, \dots, k-1$ , are known. For  $n = N-k, \dots, 0$ , and for each  $x \in \mathcal{D}_h^n$ , solve  $Y^n$ ,  $Z^n$ ,  $A^n$  and  $\Gamma^n$  by

$$(4.1) \quad \left\{ \begin{array}{l} X^{n,j} = x + b(t_n, x, Y^n, Z^n, \Gamma^n) \Delta t_{n,j} + \sigma(t_n, x, Y^n, Z^n, \Gamma^n) \Delta W_{n,j}, \\ \quad j = 1, 2, \dots, k, \\ Z^n = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} Y^{n+j} \Delta W_{n,j}^\top \right], \\ \Gamma^n = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} (Z^{n+j})^\top \Delta W_{n,j}^\top \right], \\ A^n = \sum_{j=0}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} Z^{n+j} \right], \\ -\alpha_{k,0} Y^n = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} Y^{n+j} \right] + f(t_n, x, Y^n, Z^n, \Gamma^n). \end{array} \right.$$

The main difference in Scheme 4 is that  $X^{n,j}$ ,  $Y^n, Z^n$  and  $\Gamma^n$  are all coupled together. Thus, one needs to solve nonlinear equations with some iterative procedure. In our numerical experiments, we use the following iterative scheme.

SCHEME 5. Assume random variables  $Y^{N-i}$ ,  $Z^{N-i}$  and  $\Gamma^{N-i}$  defined on  $\mathcal{D}_h^{N-i}$ ,  $i = 0, 1, \dots, k-1$ , are known.  $l$  is the iterative time,  $Y^{n,l}$ ,  $Z^{n,l}$ ,  $A^{n,l}$  and  $\Gamma^{n,l}$  are the corresponding values at the  $l$ -th iterative. For  $n = N-k, \dots, 0$ , and each  $x \in \mathcal{D}_h^n$ , solve  $Y^n$ ,  $Z^n$ ,  $A^n$  and  $\Gamma^n$  by

1. Initialize  $Y^{n,0}$ ,  $Z^{n,0}$  and  $\Gamma^{n,0}$  by  $Y^{n+1}$ ,  $Z^{n+1}$  and  $\Gamma^{n+1}$ , separately, i.e.  $Y^{n,0} = Y^{n+1}(x)$ ,  $Z^{n,0} = Z^{n+1}(x)$  and  $\Gamma^{n,0} = \Gamma^{n+1}(x)$ ;
2. Calculate  $Y^{n,l+1}$ ,  $Z^{n,l+1}$ ,  $A^{n,l+1}$  and  $\Gamma^{n,l+1}$  for  $l = 0, 1, \dots$  by

$$\left\{ \begin{array}{l} X^{n,j} = x + b(t_n, x, Y^{n,l}, Z^{n,l}, \Gamma^{n,l}) \Delta t_{n,j} + \sigma(t_n, x, Y^{n,l}, Z^{n,l}, \Gamma^{n,l}) \Delta W_{n,j}, \\ \quad j = 1, 2, \dots, k, \\ Z^{n,l+1} = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} Y^{n+j} \Delta W_{n,j}^\top \right], \\ \Gamma^{n,l+1} = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} (Z^{n+j})^\top \Delta W_{n,j}^\top \right], \\ A^{n,l+1} = \sum_{j=0}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} Z^{n+j} \right], \\ -\alpha_{k,0} Y^{n,l+1} = \sum_{j=1}^k \alpha_{k,j} \hat{\mathbb{E}}^{n,x} \left[ \mathbb{I}_{\mathcal{D}, \bar{X}^{n,j}}^{n+j} Y^{n+j} \right] + f(t_n, x, Y^{n,l+1}, Z^{n,l+1}, \Gamma^{n,l+1}). \end{array} \right.$$

Repeat the above calculation, until

$$\max \{ |Y^{n,l+1} - Y^{n,l}|, |Z^{n,l+1} - Z^{n,l}|, |A^{n,l+1} - A^{n,l}|, |\Gamma^{n,l+1} - \Gamma^{n,l}| \} < \epsilon_0;$$

3. The values  $(Y^n, Z^n, \Gamma^n, A^n)$  are set to be  $(Y^{n,l+1}, Z^{n,l+1}, A^{n,l+1}, \Gamma^{n,l+1})$ .

The computation is complete.

It is worth to remark that our Schemes 4 and 5 are heuristic generalizations of the decoupled ones in the last section. In case the drift coefficient  $b$  and the diffusion coefficient  $\sigma$  do not depend on  $Y$ ,  $Z$  and  $\Gamma$ , Scheme 5 coincides with Scheme 3.

REMARK 4.1. We have finished the construction of our multistep numerical schemes, i.e., Scheme 3 for decoupled 2FBSDEs and Scheme 5 for coupled 2FBSDEs. In both cases, the Euler method is used to discrete the forward SDEs, thus, the total computational complexity can be significantly reduced. Furthermore, we shall show by numerical tests that the solution  $(Y_t, Z_t, \Gamma_t, A_t)$  can still be of high order accuracy, in the next section.

**5. Numerical experiments.** In this section, we shall provide with several constructive numerical tests to show the efficiency and accuracy of our multistep schemes proposed in the previous sections. Applications of our numerical methods for stochastic optimal control problems will also be presented.

In all our numerical tests, we shall consider the uniform partition for the time-space domain  $[0, T] \times \mathbb{R}^m$ , for simplicity. That is, for a given  $N > 0, h > 0$ , we

set

$$\mathcal{T} := \{t_n \mid t_n = n\Delta t, \quad n = 0, 1, \dots, N, \Delta t = \frac{T}{N}\},$$

$$\mathcal{D}_h := \{x_{\mathbf{j}} \mid x_{\mathbf{j}} = x_0 + \mathbf{j} \cdot h, \quad \mathbf{j} := (j_1, j_2, \dots, j_m)^\top, \text{ for each } j_i \in \mathbf{Z}\},$$

with  $x_0 \in \mathbb{R}^m$  the initial state. Specially, for the one dimensional case, i.e.  $m = 1$ , this yields  $\mathcal{D}_h = \{x_j \mid x_j = x_0 + jh, j = 0, \pm 1, \pm 2, \dots\}$ . Our scheme is applied to calculate the values of  $Y^n$ ,  $Z^n$ ,  $\Gamma^n$  and  $A^n$ , which are the numerical values of  $Y$ ,  $Z$ ,  $\Gamma$  and  $A$  at every time-space point  $(t_n, x_{\mathbf{j}}) \in \mathcal{T} \times \mathcal{D}_h$ .

In our tests, we shall adopt the local Lagrange interpolation method  $\mathbb{I}_{\mathcal{D}}$ , so that the interpolation error estimates in (3.32) holds, and the Hermite-Gaussian quadrature rule will be used to approximate the associated conditional expectations. Since we aim to checking the convergence rates of our multistep schemes, all errors resulted from these approximations are controlled by choosing appropriate parameters (e.g. interpolation order  $r$  and the numerical of Gaussian-Hermite points). In particular, we shall use 10 Gauss-Hermite quadrature points in each dimension, and for the Lagrange interpolation, more than 6 points will be used. To balance the errors result from the time discrete truncation and the space truncation, we choose  $h = (\Delta t)^{\frac{k+1}{r+1}}$ , where  $r$  is the degree of the Langrange interpolation polynomial. Furthermore, for the  $k$ -step schemes, the information for  $\{Y^{N-j}, Z^{N-j}\}_{j=1}^k$  are needed, this can be obtained by using other numerical methods with small time steps, to maintain the high order convergence rates. Here, we just artificially assume that the values are known for simplicity.

The numerical results are obtained with FORTRAN 95 on a workstation with one Intel Xeon E5-2620 v2 CPU (12 cores, 2.10 GHz ). To accelerate the performance, OpenMP techniques are used. To guarantee the computing precision, we use the *long double* type (`real(16)`) digital for the float variables when programming. The long double variable has 34 significant digits which supplies enough precision for our computation. However, the cost of time is increasing dramatically compared with variables all defined as double (`real(8)`), but even so, the time elapsed of our programs is still competitive. In what follows, we will denote by CR the convergence rates and  $T_r$  the running time, respectively. For all our numerical examples, the terminal time  $T$  is set to be 1.0.

**5.1. Decoupled and coupled 2FBSDs.** We first test Scheme 3 for solving decoupled 2FBSDs. The first example considered is

$$(5.1) \quad \begin{cases} dX_t = \sin(t + X_t)dt + c \cos(t + X_t)dW_t, \\ -dY_t = (-\cos(t + X_t)\frac{1}{c}Z_t - \cos(t + X_t)(Y_t^2 + Y_t) - \frac{1}{4}\Gamma_t)dt - Z_t dW_t, \\ dZ_t = A_t dt + \Gamma_t dW_t, \end{cases}$$

with the initial condition  $X_{t_0} = x$  and the terminal condition  $Y_T = \sin(T + X_T)$ . It can be shown that the exact solutions are

$$Y_t = \sin(t + X_t), \quad Z_t = c \cos^2(t + X_t), \quad \Gamma_t = -2c^2 \sin(t + X_t) \cos^2(t + X_t),$$

$$A_t = -c \sin(2t + 2X_t)(1 + \sin(t + X_t)) - c^3 \cos(2t + 2X_t) \cos^2(t + X_t).$$

In the numerical test, we set  $x = 0.5$  and  $t_0 = 0$ , and solve the 2FBSDs by scheme 3 with different step parameter  $k$ . The numerical errors  $|Y^0 - Y_0|$ ,  $|Z^0 - Z_0|$ ,  $|\Gamma^0 - \Gamma_0|$

and  $|A^0 - A_0|$  and the corresponding convergence rates are listed in Table 5.1. It is shown in Table 5.1 that: the multistep numerical scheme works very well, and the numerical error goes to machine accuracy. Moreover, the method admits a  $k$ -order convergence rate, and it remains stable for  $1 \leq k \leq 6$ , which is coincide with the classic numerical ODEs theory and our previous results [26].

Furthermore, given a fixed accuracy tolerance, it would be more efficient if the multistep scheme with a large  $k$  is used. This can be more easily seen from the following table.

STEP	N	$T_{rn}$	$ Y^0 - Y_0 $	$ Z^0 - Z_0 $	$ \Gamma^0 - \Gamma_0 $	$ A^0 - A_0 $
$k = 1$	8192	135.0s	5.172E-05	2.739E-05	7.573E-07	3.585E-07
$k = 2$	2048	10.82s	1.232E-09	4.704E-08	2.301E-08	1.920E-07
$k = 3$	512	5.58s	3.423E-09	3.097E-09	7.797E-10	4.677E-09
$k = 4$	128	2.85s	1.159E-09	1.733E-09	2.113E-09	6.042E-09

We now consider an example with the geometry Brownian motion in the forward SDE, the main feature here is that the drift and diffusion terms are unbounded. The example yields

$$(5.2) \quad \begin{cases} dX_t = rX_t dt + cX_t dW_t, \\ -dY_t = -e^{-\frac{x_t^2}{M}} + \frac{2}{M}rX_t^2 Y_t - \frac{\Gamma_t}{2} + \frac{c}{2}Z_t dt - Z_t dW_t, \\ dZ_t = A_t dt + \Gamma_t dW_t \end{cases}$$

with terminal condition  $Y_T = Te^{-\frac{x_T^2}{M}}$ . It can be shown that the exact solution is

$$\begin{aligned} Y_t &= te^{-\frac{x_t^2}{M}}, \quad Z_t = -\frac{2c}{M}tX_t^2 e^{-\frac{x_t^2}{M}}, \quad \Gamma_t = \frac{4c^2}{M^2}tX_t^2 e^{-\frac{x_t^2}{M}}(X_t^2 - M), \\ A_t &= -\frac{2c}{M}e^{-\frac{x_t^2}{M}}[(1 + 2rt + c^2t)X_t^2 - \frac{2r + c^2}{M}tX_t^4 + \frac{2c^2t}{M}X_t^6]. \end{aligned}$$

In this example, we set  $x_0 = 1.5$ , and the parameters are chosen as  $r = 0.2, c = 0.01, M = 4$ . The corresponding numerical results are listed the in Table 5.2. Again, our multistep schemes behaves very well, and high order convergence rates are obtained.

We now test Scheme 5 for solving coupled 2FBSDEs. Here, iterative process is needed. To show the high accuracy of the proposed schemes, we set the tolerance as  $\epsilon = 10^{-25}$ . We first consider the following coupled 2FBSDEs:

$$(5.3) \quad \begin{cases} dX_t = (\sin(t + X_t) + Z_t/c + \sin(t + X_t)Y_t - 1)dt \\ \quad + (c \cos(t + X_t) - c + c \cos^2(t + X_t) + c \sin(t + X_t)Y_t)dW_t, \\ -dY_t = -\cos(t + X_t)\frac{1}{c}Z_t - \cos(t + X_t)(Y_t^2 + Y_t) - \frac{1}{4}\Gamma_t dt - Z_t dW_t, \\ dZ_t = A_t dt + \Gamma_t dW_t, \end{cases}$$

with conditions  $Y_T = \sin(T + X_T)$  and  $X_{t_0} = x$ . The exact solution is

$$\begin{aligned} Y_t &= \sin(t + X_t), \quad Z_t = c \cos^2(t + X_t), \quad \Gamma_t = -2c^2 \sin(t + X_t) \cos^2(t + X_t), \\ A_t &= -c \sin(2t + 2X_t)(1 + \sin(t + X_t)) - c^3 \cos(2t + 2X_t) \cos^2(t + X_t). \end{aligned}$$

TABLE 5.1  
Numerical results of example 1 with  $c = 0.1$

STEP		N	$ Y^0 - Y_0 $	$ Z^0 - Z_0 $	$ \Gamma^0 - \Gamma_0 $	$ A^0 - A_0 $	$T_{rn}$
K=1	DT	32	1.253E-02	6.800E-03	3.156E-04	1.703E-03	1.87s
		64	6.411E-03	3.466E-03	1.854E-04	1.059E-03	
		128	3.248E-03	1.752E-03	8.667E-05	4.409E-04	
		256	1.637E-03	8.808E-04	3.570E-05	1.381E-04	
		512	8.227E-04	4.405E-04	1.411E-05	3.472E-05	
	CR		0.98	0.99	1.13	1.42	
K=2	DT	32	6.123E-06	1.897E-04	1.086E-04	9.312E-04	4.96s
		64	1.334E-06	4.761E-05	2.614E-05	2.210E-04	
		128	3.332E-07	1.192E-05	6.207E-06	5.196E-05	
		256	8.494E-08	2.995E-06	1.500E-06	1.253E-05	
		512	2.133E-08	7.514E-07	3.707E-07	3.098E-06	
	CR		2.03	2.00	2.05	2.06	
K=3	DT	32	1.149E-05	1.409E-05	2.846E-06	1.988E-05	13.96s
		64	1.615E-06	1.668E-06	3.734E-07	2.382E-06	
		128	2.117E-07	2.026E-07	4.848E-08	2.972E-07	
		256	2.704E-08	2.496E-08	6.202E-09	3.727E-08	
		512	3.415E-09	3.096E-09	7.818E-10	4.650E-09	
	CR		2.93	3.04	2.96	3.01	
K=4	DT	32	2.853E-07	4.169E-07	6.203E-07	2.171E-06	17.74s
		64	1.843E-08	2.700E-08	3.588E-08	1.122E-07	
		128	1.161E-09	1.736E-09	2.097E-09	5.886E-09	
		256	7.271E-11	1.106E-10	1.275E-10	3.463E-10	
		512	4.545E-12	6.973E-12	7.852E-12	2.097E-11	
	CR		3.99	3.97	4.07	4.17	
K=5	DT	32	1.124E-08	3.151E-08	2.563E-08	1.660E-07	34.41s
		64	5.297E-10	9.433E-10	8.493E-10	4.027E-09	
		128	1.877E-11	2.770E-11	2.632E-11	9.995E-11	
		256	6.180E-13	8.174E-13	8.310E-13	2.951E-12	
		512	1.971E-14	2.480E-14	2.693E-14	9.838E-14	
	CR		4.80	5.07	4.97	5.18	
K=6	DT	32	4.721E-10	9.254E-10	5.225E-09	1.144E-08	50.75s
		64	7.847E-12	1.542E-11	7.602E-11	1.545E-10	
		128	1.144E-13	2.506E-13	1.140E-12	2.144E-12	
		256	1.741E-15	4.074E-15	1.710E-14	2.863E-14	
		512	2.674E-17	6.557E-17	2.622E-16	4.232E-16	
	CR		6.03	5.94	6.06	6.18	
K=7	DT	32	1.422E-11	1.805E-11	4.311E-10	7.197E-09	83.65s
		64	3.876E-13	8.371E-13	1.417E-12	3.941E-12	
		128	2.365E-15	1.959E-15	2.089E-14	3.823E-13	
		256	2.762E-17	1.366E-17	1.265E-15	1.528E-14	
		512	4.993E-18	1.958E-16	9.029E-17	1.894E-16	
	CR		5.67	4.89	5.45	5.84	
K=8	DT	32	7.015E-12	1.501E-10	2.182E-09	2.147E-08	108.42s
		64	4.860E-12	6.270E-11	3.480E-10	7.047E-09	
		128	1.117E-10	6.105E-09	8.660E-08	1.014E-06	
		256	7.367E-03	3.798E+00	4.029E+02	3.354E+04	
		512	NaN	NaN	NaN	NaN	
	CR		NaN	NaN	NaN	NaN	

The numerical results are shown in Table 5.3, where one can concludes that the numerical methods work well and high-order convergence rates are obtained. However, due to the computational complexity, numerical results with only multistep up to  $k = 3$  are presented.



Our next couple 2FBSDEs example yields

$$\left\{ \begin{array}{l} X_t = x + \int_0^t \frac{1}{(1 + \exp(s + X_s))(1 + Y_s)} ds + \int_0^t Y_s dW_s, \\ Y_t = \frac{\exp(T + X_T)}{1 + \exp(T + X_T)} \\ \quad + \int_t^T \left[ \frac{2Y_s}{1 + 2\exp(s + X_s)} + \frac{1}{2} \left( \Gamma_s - \frac{Y_s Z_s}{1 + \exp(s + X_s)} \right) \right] ds - \int_t^T Z_s dW_s, \\ Z_t = Z_T - \int_t^T A_s ds - \int_t^T \Gamma_s dW_s. \end{array} \right.$$

TABLE 5.2  
Numerical results of example 2.

STEP		N	$ Y^0 - Y_0 $	$ Z^0 - Z_0 $	$ \Gamma^0 - \Gamma_0 $	$ A^0 - A_0 $	$T_{rn}$
K=1	DT	32	4.808E-03	2.331E-04	2.431E-06	6.699E-07	2.58s
		64	2.402E-03	1.169E-04	1.232E-06	3.335E-07	
		128	1.200E-03	5.854E-05	6.207E-07	1.665E-07	
		256	5.999E-04	2.929E-05	3.115E-07	8.308E-08	
		512	2.999E-04	1.465E-05	1.560E-07	4.143E-08	
	CR		1.00	1.00	0.99	1.00	
K=2	DT	32	1.385E-05	1.294E-06	1.459E-07	1.457E-06	7.54s
		64	3.582E-06	3.252E-07	3.653E-08	3.603E-07	
		128	9.098E-07	8.147E-08	9.140E-09	8.954E-08	
		256	2.291E-07	2.039E-08	2.286E-09	2.231E-08	
		512	5.747E-08	5.098E-09	5.717E-10	5.569E-09	
	CR		1.98	2.00	2.00	2.01	
K=3	DT	32	3.490E-07	5.124E-08	1.213E-09	8.537E-09	16.73s
		64	4.638E-08	6.484E-09	1.552E-10	1.130E-09	
		128	5.983E-09	8.157E-10	1.960E-11	1.454E-10	
		256	7.604E-10	1.023E-10	2.461E-12	1.846E-11	
		512	9.590E-11	1.281E-11	3.082E-13	2.326E-12	
	CR		2.96	2.99	2.99	2.96	
K=4	DT	32	4.889E-09	8.085E-10	7.806E-11	6.895E-10	34.39s
		64	3.369E-10	5.171E-11	4.934E-12	4.232E-11	
		128	2.212E-11	3.276E-12	3.100E-13	2.601E-12	
		256	1.417E-12	2.063E-13	1.942E-14	1.610E-13	
		512	8.969E-14	1.295E-14	1.215E-15	1.003E-14	
	CR		3.94	3.98	3.99	4.02	
K=5	DT	32	4.916E-11	7.997E-12	6.357E-14	2.573E-12	61.72s
		64	1.548E-12	2.514E-13	1.625E-15	7.997E-14	
		128	4.545E-14	7.896E-15	3.486E-17	2.277E-15	
		256	1.316E-15	2.481E-16	6.573E-19	6.360E-17	
		512	3.843E-17	7.792E-18	1.067E-20	1.790E-18	
	CR		5.08	4.99	5.63	5.12	
K=6	DT	32	2.246E-12	3.449E-13	3.677E-14	4.326E-13	93.42s
		64	3.805E-14	5.534E-15	5.708E-16	6.289E-15	
		128	5.870E-16	8.807E-17	9.179E-18	9.283E-17	
		256	8.808E-18	1.394E-18	1.463E-19	1.391E-18	
		512	1.319E-19	2.197E-20	2.318E-21	2.136E-20	
	CR		6.01	5.98	5.98	6.07	
K=7	DT	32	3.908E-13	1.027E-13	2.830E-15	1.527E-13	196.54s
		64	4.974E-16	1.217E-15	3.787E-17	1.929E-15	
		128	3.750E-17	1.204E-17	3.601E-18	1.489E-16	
		256	3.893E-19	2.617E-17	2.460E-15	3.265E-13	
		512	5.287E-17	8.003E-15	6.419E-12	1.677E-09	
	CR		3.60	1.29	-2.83	-3.42	

TABLE 5.3  
Numerical results of example 3.

STEP		N	$ Y^0 - Y_0 $	$ Z^0 - Z_0 $	$ \Gamma^0 - \Gamma_0 $	$ A^0 - A_0 $	$T_{rn}$
K=1	DT	32	4.126E-02	9.266E-03	3.980E-03	5.001E-02	35.42s
		64	2.085E-02	4.782E-03	1.996E-03	2.532E-02	
		128	1.048E-02	2.430E-03	9.968E-04	1.273E-02	
		256	5.255E-03	1.225E-03	4.977E-04	6.377E-03	
		512	2.631E-03	6.149E-04	2.486E-04	3.192E-03	
	CR		0.99	0.98	1.00	0.99	
K=2	DT	32	5.077E-04	4.045E-05	2.561E-04	1.640E-03	232.23s
		64	1.250E-04	7.443E-06	6.648E-05	4.266E-04	
		128	3.099E-05	1.541E-06	1.690E-05	1.086E-04	
		256	7.713E-06	3.462E-07	4.258E-06	2.740E-05	
		512	1.924E-06	8.166E-08	1.069E-06	6.876E-06	
	CR		2.01	2.23	1.98	1.98	
K=3	DT	32	6.995E-05	2.551E-05	2.388E-05	1.568E-04	987s
		64	8.908E-06	3.156E-06	2.693E-06	1.840E-05	
		128	1.123E-06	3.923E-07	3.173E-07	2.225E-06	
		256	1.409E-07	4.889E-08	3.843E-08	2.733E-07	
		512	1.764E-08	6.103E-09	4.726E-09	3.383E-08	
	CR		2.99	3.01	3.07	3.04	

The associated exact solution is

$$\begin{aligned}
Y_t &= \frac{\exp(t + X_t)}{1 + \exp(t + X_t)}, \quad Z_t = \frac{\exp(t + X_t)^2}{(1 + \exp(t + X_t))^3}, \\
\Gamma_t &= \frac{\exp(t + X_t)^3(2 - \exp(t + X_t))}{(1 + \exp(t + X_t))^5}, \\
A_t &= \frac{2 \exp(t + X_t)^2(2 - \exp(t + X_t))}{(1 + \exp(t + X_t))^3(1 + 2 \exp(t + X_t))} \\
&\quad + \frac{\exp(t + X_t)^4(\exp(t + X_t)^2 - 7 \exp(t + X_t) + 4)}{2(1 + \exp(t + X_t))^7}.
\end{aligned}$$

The corresponding numerical results are listed in Table 5.4. Similar convergence results are shown as the above numerical tests.

**5.2. Applications to stochastic optimal control.** We now show that one can solve stochastic optimal control problems in a 2FBSDEs way by using our multistep numerical schemes. To this end, let us consider the control problem, whose dynamic state equation is described by a forward SDE

$$(5.4) \quad dX_t = b(t, X_t, \alpha_t)dt + \sigma(t, X_t, \alpha_t)dW_t,$$

with the cost functional

$$(5.5) \quad J(\alpha) = \mathbb{E} \left[ \int_0^T f(t, X_t, \alpha_t)dt + g(X_T) \right].$$

The goal is to minimize the cost functional, i.e. Find  $\alpha^* \in \mathcal{U}$  (where  $\mathcal{U}$  contains all admissible controls, see e.g. [16] for the corresponding definition) satisfying

$$V(t, x) := J(t, x; \alpha^*) = \inf_{\alpha \in \mathcal{U}} J(t, x; \alpha).$$

TABLE 5.4  
Numerical results of example 4.

STEP		N	$ Y^0 - Y_0 $	$ Z^0 - Z_0 $	$ \Gamma^0 - \Gamma_0 $	$ A^0 - A_0 $	$T_{rn}$
K=1	DT	32	1.079E-03	2.877E-03	3.847E-03	3.427E-03	21.54s
		64	5.374E-04	1.463E-03	2.008E-03	1.796E-03	
		128	2.680E-04	7.389E-04	1.030E-03	9.216E-04	
		256	1.338E-04	3.715E-04	5.225E-04	4.678E-04	
		512	6.681E-05	1.864E-04	2.634E-04	2.359E-04	
	CR		1.00	0.99	0.97	0.97	
K=2	DT	32	5.131E-05	1.750E-05	2.281E-04	3.383E-04	70.10s
		64	1.328E-05	4.871E-06	5.814E-05	8.806E-05	
		128	3.388E-06	1.248E-06	1.477E-05	2.276E-05	
		256	8.523E-07	3.171E-07	3.691E-06	5.698E-06	
		512	2.139E-07	7.953E-08	9.282E-07	1.432E-06	
	CR		1.98	1.95	1.99	1.97	
K=3	DT	32	3.134E-06	6.034E-06	4.653E-06	2.167E-05	259.11s
		64	4.058E-07	7.806E-07	5.391E-07	2.879E-06	
		128	5.121E-08	9.992E-08	5.827E-08	3.605E-07	
		256	6.407E-09	1.257E-08	7.059E-09	4.530E-08	
		512	8.052E-10	1.574E-09	8.454E-10	5.398E-09	
	CR		2.98	2.98	3.11	2.99	

One can construct the corresponding HJB equation as

$$\frac{\partial}{\partial t} V(t, x) + \inf_{\alpha \in \mathcal{U}} \left\{ \frac{\sigma(t, x, \alpha)^2}{2} \frac{\partial^2}{\partial x^2} V(t, x) + b(t, x, \alpha) \frac{\partial}{\partial x} V(t, x) - f(t, x, \alpha) \right\} = 0.$$

and moreover, we have

$$(5.6) \quad \alpha^*(t, x) = \arg \inf_{\alpha \in \mathcal{U}} \left\{ \frac{\sigma(t, x, \alpha)^2}{2} \frac{\partial^2 V}{\partial x^2} + b(t, x, \alpha) \frac{\partial V}{\partial x} - f(t, x, \alpha) \right\}.$$

By inserting (5.6) into the HJB equation, one shows that the cost function satisfies

$$(5.7) \quad \frac{\partial}{\partial t} V(t, x) + G(t, x, \frac{\partial}{\partial x} V(t, x), \frac{\partial^2}{\partial x^2} V(t, x)) = 0,$$

where

$$G(t, x, p, P) = \frac{\sigma(t, x, \alpha^*)^2}{2} P + b(t, x, \alpha^*) p - f(t, x, \alpha^*).$$

For the above nonlinear PDEs, one can construct a corresponding 2FBSDEs. Then, one can solve the associate 2FBSDEs using our numerical schemes to obtain  $(X_t, Y_t, Z_t, \Gamma_t)$ , which yields the optimal state for the control problem, and finally, we can get obtain in view of (5.6) that

$$(5.8) \quad \alpha_t^* = g(X_t, Y_t, Z_t, \Gamma_t)$$

with  $g(\cdot)$  being certain functions. This is a new approach dealing with the optimal control problem in a 2FBSDEs way. Now, we illustrate the idea by the following example.

**Tracking a particle under the microscope:** consider the following system

$$(5.9) \quad dX_t = \beta \alpha_t dt + \sigma dW_t,$$

where  $X_t$  is the distance between the particle and the focus of the microscope,  $\beta \in \mathbb{R}$  is the gain in our servo loop and  $\sigma > 0$  is the diffusion constant of the particle. We

would like to keep the particle in focus, i.e. we expect that  $X_t$  is as close to zero as possible. However, we have to introduce a power constraint on the control as well, as we cannot drive the servo motor with arbitrarily large input powers. We thus introduce the control cost

$$J(\alpha) = \mathbb{E} \left[ p \int_0^T X_t^2 dt + q \int_0^T \alpha^2 dt \right],$$

where  $p, q > 0$  allows us to select the tradeoff between good tracking and low feedback power. One can construct the associate bellman equation

$$\begin{aligned} 0 &= \frac{\partial}{\partial t} V(t, x) + \inf_{\alpha \in \mathbb{R}} \left\{ \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} V(t, x) + \beta \alpha \frac{\partial}{\partial x} V(t, x) + p x^2 + q \alpha^2 \right\} \\ &= \frac{\partial}{\partial t} V(t, x) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} V(t, x) - \frac{\beta}{4q} \left( \frac{\partial}{\partial x} V(t, x) \right)^2 + p x^2 \end{aligned}$$

with  $V(T, x) = 0$ . Furthermore, one can show that the optimal control parameter (i.e., exact solution) is

$$(5.10) \quad \alpha_t^* = -\frac{\beta}{2q} \frac{\partial}{\partial x} V(t, x).$$

The classic numerical solution would relies on solving the above bellman equation.

We now solve the problem in a 2FBSDEs way, to this end, we first construct the corresponding 2FBSDEs as follows

$$(5.11) \quad \begin{cases} dX_t = \beta c dt + \sigma dW_t, \\ -dY_t = \left( -\frac{\beta^2}{4q\sigma^2} Z_t^2 - \frac{\beta c}{\sigma} Z_t + p X_t^2 \right) dt - Z_t dW_t, \\ dZ_t = A_t dt + \Gamma_t dW_t. \end{cases}$$

By solving this 2FBSDEs, we obtain the numerical solution in view of (5.10) by

$$(5.12) \quad \alpha^n = \frac{-\beta}{2q\sigma} Z^n.$$

In the numerical test, we set  $\mu = 0.1$ ,  $r = 0.03$ ,  $\sigma = 0.5$ ,  $c = 0.1$ . The numerical results are shown in Table 5.5. It can be seen from Table 5.5 that the approach is of high order accuracy, both for the 2FBSDE solution and the optimal control  $\alpha$ .

**6. Conclusions.** We have extended our multistep schemes in [26] to the multistep schemes 3 and 5 for solving the second order FBSDEs. The key feature of the proposed multistep schemes is that the Euler method is used to discrete the forward SDE, which dramatically reduces the entire computational complexity. Furthermore, it is shown that the quantities of interest (e.g., the solution tuple  $(Y_t, Z_t, A_t, \Gamma_t)$  in the 2FBSDEs) are still of high order accuracy. Several numerical examples are presented to show the effective of the proposed numerical schemes. Applications of our numerical schemes for stochastic optimal control problems are also discussed.

There are, however, some other related topics that need to be investigated:

- High dimensional problems. Note the methods here can be easily extended to high dimensional problems. However, we have proposed the local Lagrange

TABLE 5.5  
Numerical results of stochastic optimal control.

STEP		N	$ Y^0 - Y_0 $	$ Z^0 - Z_0 $	$ \Gamma^0 - \Gamma_0 $	$ \alpha^0 - \alpha^* $	$T_{rn}$
K=1	DT	32	5.647E-03	5.856E-02	1.420E-02	1.952E-02	2.05s
		64	2.788E-03	2.912E-02	7.075E-03	9.706E-03	
		128	1.385E-03	1.452E-02	3.531E-03	4.839E-03	
		256	6.903E-04	7.249E-03	1.764E-03	2.416E-03	
		512	3.446E-04	3.622E-03	8.814E-04	1.207E-03	
	CR		1.01	1.00	1.00	1.00	
K=2	DT	32	1.438E-03	8.777E-04	5.823E-05	2.926E-04	6.74s
		64	3.660E-04	2.192E-04	1.471E-05	7.308E-05	
		128	9.230E-05	5.478E-05	3.696E-06	1.826E-05	
		256	2.317E-05	1.369E-05	9.264E-07	4.564E-06	
		512	5.806E-06	3.422E-06	2.319E-07	1.141E-06	
	CR		1.99	2.00	1.99	2.00	
K=3	DT	32	3.439E-06	4.495E-06	1.848E-06	1.498E-06	19.44s
		64	4.758E-07	5.315E-07	2.276E-07	1.772E-07	
		128	6.243E-08	6.456E-08	2.824E-08	2.152E-08	
		256	7.992E-09	7.953E-09	3.517E-09	2.651E-09	
		512	1.011E-09	9.869E-10	4.387E-10	3.290E-10	
	CR		2.94	3.04	3.01	3.04	
K=4	DT	32	4.583E-07	3.240E-07	3.570E-08	1.080E-07	57.62s
		64	2.986E-08	2.005E-08	2.264E-09	6.682E-09	
		128	1.902E-09	1.246E-09	1.425E-10	4.154E-10	
		256	1.200E-10	7.767E-11	8.933E-12	2.589E-11	
		512	7.535E-12	4.847E-12	5.592E-13	1.616E-12	
	CR		3.97	4.01	3.99	4.01	
K=5	DT	32	4.302E-09	2.299E-10	8.347E-10	7.662E-11	94.61s
		64	1.512E-10	1.904E-11	2.438E-11	6.346E-12	
		128	4.996E-12	7.749E-13	7.352E-13	2.583E-13	
		256	1.604E-13	2.698E-14	2.256E-14	8.995E-15	
		512	5.080E-15	8.862E-16	6.985E-16	2.954E-16	
	CR		4.93	4.54	5.05	4.54	

interpolation methods here in our schemes. For high dimensional problems, this would results in the tensor Lagrange interpolation methods, which may be time consuming. Thus, we would suggest more feasible techniques such as the sparse grid interpolation, RBF interpolation, etc. This would be part of our future studies.

- Rigorous stability and convergence analysis. This is also our ongoing project.

#### REFERENCES

- [1] C. Bender and J. Zhang, Time discretization and Markovian iteration for coupled FBSDEs, *Ann. Appl. Probab.*, 18(2008), pp. 143-177.
- [2] J. M. Bismut, Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.*, 44(1973), pp. 384-404.
- [3] B. Bouchard and N. Touzi, Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations, *Stochastic Process. Appl.*, 111(2004), pp. 175-206.
- [4] J.F. Chassagneux and D. Crisen, Runge-Kutta schemes for BSDEs, to appear in *Ann. Appl. Probab.*, 2014.
- [5] P. Cheridito, H. M. Soner, N. Touuzi, and Nicolas Victoir, Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs, *Communications on Pure and Applied Mathematics*, Vol. LX (2007), pp. 1081-1110.
- [6] D. Crisen and K. Manolarakis, Solving backward stochastic differential equations using the cubature method, *SIAM J. Math. Finance*, (3)2012, pp. 534-571.
- [7] J. Douglas, J. Ma and P. Protter, Numerical methods for forward-backward stochastic differential equations, *Ann. Appl. Probab.*, 6(1996), pp. 940-968.

- [8] Arash Fahim, Nizar Touzi, and Xavier Warin, A probabilistic numerical method for fully non-linear parabolic PDEs, *Ann. Appl. Probab.*, 4(2011), pp. 1322-1364.
- [9] Y. Fu, W. Zhao, and T. Zhou, Efficient sparse grid approximations for multi-dimensional coupled forward backward stochastic differential equations, submitted, 2015.
- [10] W. Guo, J. Zhang, and J. Zhuo, A Monotone Scheme for High Dimensional Fully Nonlinear PDEs, arXiv:1212.0466, to appear in *Ann. Appl. Probab.*, 2015.
- [11] N. El Karoui, C. Kapoudjian, E. Pardoux, S. G. Peng and M. C. Quenez, Reflected solutions of backward SDE's and related obstacle problems for PDE's, *Ann. Probab.*, 25(1997), pp. 702-737.
- [12] N. EL Karoui, S. G. Peng, and M. C. Quenez, Backward stochastic differential equations in finance, *Math. Finance*, 7(1997), pp. 1-71.
- [13] J. P. Lemor, E. Gobet and X. Warin, A regression-based Monte Carlo method for backward stochastic differential equations, *Ann. Appl. Probab.*, 15(2005), pp. 2172-2202.
- [14] J. Ma P. Protter and J. Yong, Solving forward-backward stochastic differential equations explicitly - a four step scheme, *Probab. Theory Related Fields*, 98(1994), pp. 339-359.
- [15] J. Ma J. Shen and Y. Zhao, On numerical approximations of forward-backward stochastic differential equations, *SIAM J. Numer. Anal.*, 46 (2008), pp. 2636-2661.
- [16] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Mathematics, vol. 1702. Berlin: Springer.
- [17] G. N. Milstein and M. V. Tretyakov, Numerical algorithms for forward-backward stochastic differential equations, *SIAM J. Sci. Comput.*, 28(2006), pp. 561-582.
- [18] B. Oksendal, *Stochastic Differential Equations*, Six Edition, Springer-Verlag, Berlin, 2003.
- [19] E. Pardoux and S. G. Peng, Adapted solution of a backward stochastic differential equation, *Systems Control Lett.*, 14(1990), pp. 55-61.
- [20] E. Pardoux and S. Tang, Forward-backward stochastic differential equations and quasilinear parabolic PDEs, *Probab. Theory Relat. Fields*, 114(1999), pp. 123-150.
- [21] S. G. Peng, Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, *Stoch. Stoch. Repts.*, 37 (1991), pp. 61-74.
- [22] H. M. Soner, N. Touzi, and J. Zhang, Wellposedness of second order backward SDEs, *Probab. Theory Relat. Fields*, Vol. 153(2012), pp:149-190.
- [23] T. Tang, W. Zhao, and T. Zhou, Deferred correction methods for forward backward stochastic differential equations, submitted, 2015.
- [24] J. Zhang, A numerical scheme for BSDEs, *Ann. Appl. Probab.*, 14(2004), pp. 459-488.
- [25] W. Zhao, L. Chen and S. Peng, A new kind of accurate numerical method for backward stochastic differential equations, *SIAM J. Sci. Comput.*, 28(2006), pp. 1563-1581.
- [26] W. Zhao, Y. Fu, and T. Zhou, New kinds of high-order multistep schemes for coupled forward backward stochastic differential equations, *SIAM J. Sci. Comput.*, 36(4), pp. A1731-1751, 2014.
- [27] W. Zhao, G. Zhang and L. Ju, A stable multistep scheme for solving backward stochastic differential equations, *SIAM J. Numer. Anal.*, 48(2010), pp. 1369-1394.
- [28] W. Zhao, W. Zhang and L. Ju, A numerical method and its error estimates for the decoupled forward-backward stochastic differential equations, *Commun. Comput. Phys.*, 15(2014), pp. 618-646.